Analysis of Customer Relationship Management based on Parameter Estimation of Expectation-Maximization Reliability Model

Tiantian YANG

Department of E-Commerce, Jinhua Polytechnic, Jinhua, Zhejiang, 321000, China

Abstract — Help is needed by enterprises to effectively improve their customer relationship by implementing customer relationship management to gain the competitive advantage in the market. To deal with this we propose a method of customer relationship management, evaluation and analysis based on the parameter estimation of expectation-maximization reliability model. Firstly, we construct an evaluation system of corporate customer relationship by structural equation modeling. Secondly, we build an expectation-maximization mathematical expression form for generalized grouped data by using the characteristic of independent vector distribution in the characteristics of Poisson distribution. We realize the maximum estimation of log-likelihood function (ML) of model parameters during the E-step and M-step processes of model estimation. Finally, we verify the effectiveness of the algorithm through experiment.

Keywords— expectation maximization; parameter estimation; customer relationship; management evaluation; poisson distribution

I. INTRODUCTION

The customer of enterprise is often called “Employer” or “Party A”, and a good relationship between enterprise and employer can promote the improvement of project performance. It has been proved that the “Partnering” mode widely used in world is an effective method. However, what’s more important is the ability to build and manage such complex relationship well, namely customer relationship ability. By introducing the customer relationship management (CRM), enterprises can avoid the risk from customers and improve the customer satisfaction to realize a win-win result [1]. Therefore, how to objectively and scientifically evaluate the corporate customer relationship ability is significant for enterprises to set their development objectives, optimize their resource allocation and improve their economic benefit.

The evaluation of corporate customer relationship ability is a complex work, which has some requirements on the constitution of evaluation indexes and the selection of evaluation methods. The existing researches on the customer relationship ability are relatively limited, and relevant empirical researches are more lacking. Joseph et al [2] put forward the core ability of EPC (Engineer Procure Construct, Design-Purchase-Construction) enterprises, including relationship ability; Qiu Huifang et al [3] constructed the system of factors in customer relationship ability by normative analysis. Currently, in the methods of enterprise ability evaluation, the Analytic Hierarchy Process (AHP), Fuzzy Comprehensive Evaluation (FCE) and Data Envelopment Analysis (DEA) often ignore the possible mutual effect between evaluation indexes, but Structural Equation Modeling (SEM) can process the relationship between multiple variables simultaneously, estimate the factor structure and factor relationship simultaneously, allow the measuring errors in independent variables and dependent variables, and permit the measurement model with a higher elasticity. On account of these, using the SEM, the constitution of customer relationship ability evaluation system is discussed in the paper to provide a reference for effectively improve the customer relationship ability.

Aiming at the problem about customer relationship management, an enhanced parameter estimation based on expectation-maximization reliability model is proposed in the paper. It constructs an evaluation system of corporate customer relationship ability by Structural Equation Modeling, builds an expectation-maximization mathematical expression form for generalized grouped data by using the characteristic of independent vector distribution in the characteristics of Poisson distribution, and verifies the effectiveness of such algorithm through experiment.

II. EVALUATION SYSTEM OF CORPORATE CUSTOMER RELATIONSHIP ABILITY

Basing on resource based view and enterprise ability theory, the evaluation system of corporate customer relationship ability constructed in the paper includes 6 dimensionality:

(1) Support ability of information technology: In the information age, the customer relationship management of enterprise cannot be separated from the support of information technology. At present, telephone, fax, E-mail, company homepage and relatively simple database are the information technologies commonly used by Chinese enterprises. For example, some large-scale Chinese enterprises set “Interactive Communication” modules on the homepages of their companies, and set “Complaint and Consultation, Service Hotline, Business Platform, Online Massage and Online Survey” below that, in which the Business Platform including CRM Membership System, File Management System and Project Management System is conducive to improve the management level of customer relationship.
(2) Support ability of normative system: Guarantee the corporate customer relationship management can be implemented orderly by developing following normative documents: 1) Process, such as management process of bidding and tendering, describing the steps of bidding work and the roles and responsibilities of enterprise; 2) System, such as project warranty system, namely determining the specific works shall be done and the standard shall be reached for project warranty system on the basis of process; 3) Model, such as model of application for contract section commencement; enterprises only need filling the blank model before commencement, and the work efficiency can be improved in this way.

(3) Support ability of organization system, including organization structure design and personnel allocation: Design a reasonable organization structure, integrate each functional department, such as market bidding department, project management department and quality assurance department, contribute to improve the customer relationship ability of enterprises. Staffs are the subjects in the activity of customer relationship management. It is very important to allocate staffs with professional skills and plentiful experiences.

(4) Communication ability of customer relationship: The construction industry is an information-intensive industry, with a lot of project information sharing between enterprises and employers. The communication is not only related to the realization degree of project objectives, but also is an important constitution of customer relationship ability for enterprises [4, 5].

(5) Cooperative ability of customer relationship: Due to the huge investment, long production period and the indeterminacy of construction products, the cooperation between enterprises and employers cannot be ignored. The cooperation between organizations is related to the relationship ability. It is shown in surveys that employers think cooperation is the most important, which means enterprises shall pay full attention to cooperation.

(6) Coordination ability of customer relationship: Enterprises and employers are interdependent, so the coordination between them is indispensable. Coordinating with employers is a necessary skill for enterprises to develop the customer relationship ability. The coordination methods commonly adopted are site meeting, third-party mediation and lawsuit.

III. ENHANCED PARAMETER ESTIMATION BASED ON EXPECTATION-MAXIMIZATION RELIABILITY MODEL

A. Expectation-Maximization Algorithm

Maximum likelihood (ML) is commonly used to obtain the SRM model parameters based on NHPP. In the case of maximum likelihood, the parameter estimation is realized by maximizing the log-likelihood function (LLF) of given parameter. Traditionally, the parameters are generally estimated by two kinds of data:

(1) Time data: a set of data about customer relationship management observed on the basis of time, which called time data (of customer relationship management), giving a serial accurate CRM time. Specifically, time data can be represented as \( T = \{t_1, t_2, \ldots, t_n\} \), where \( t_i \) is the \( i \)th CRM time, \( m_t \) is the total number of customer relationship detected during the observation time of \( t \), namely \( M(t_i) = m_t \). When the observation time coincides with CRM time \( t_i = t_m \), the LLF process of the time data of SRM model based on NHPP can be obtained.

\[
LLF(\omega, \theta | T) = \sum_{i=1}^{n} \log(\omega f(t_i; \theta)) - \omega F(t_i; \theta)
\]  
(1)

where \( \theta \) is the CRM time distribution of a parameter vector. The calculation form of \( f(t; \theta) \) shall be:

\[
f(t; \theta) = \frac{dF(t; \theta)}{dt}
\]  
(2)

(2) Grouped data: the grouped data is formed by the total number of customer relationship detected in given time interval. The grouped data can be represented as \( G = [(x_i, y_i), \ldots, (x_n, y_n)] \), where \( y_i \) is the cumulative number of customer relationship in the time interval of \([0, x_i]\) . The LLF process of grouped data of SRM model based on NHPP is:

\[
LLF(\omega, \theta | G) = \sum_{i=1}^{n} \log(y_i - y_{i-1}) \log(\omega F(x_i; \theta) - F(x_i; \theta))
\]

\[
- \omega F(x_i; \theta) - \sum_{i=1}^{n} \log(y_i - y_{i-1})!
\]  
(3)

Where \( x_0 = 0 \), \( y_0 = 0 \).

(3) Generalized grouped data: supposed that the form of generalized grouped data is \( l = [(x_i, u_i, z_i), \ldots, (x_n, u_n, z_n)] \), where \( u_i \) is the number of customer relationship detected in the section of \( (x_{i-1}, x_i) \); \( z_i \) is a indicator variable, representing the number of customer relation detected in the moment of \( x_i \). For generalized grouped data, the LLF process of grouped data of SRM model based on NHPP is:

\[
LLF(\omega, \theta | I) = \sum_{i=1}^{n} u_i \log(\omega F(x_i; \theta) - F(x_i; \theta))
\]

\[
+ \sum_{i=1}^{n} z_i \log(\omega f(x_i; \theta)) - \omega F(x_i; \theta) - \sum_{i=1}^{n} \log(u_i)! \]  
(4)

Expectation-maximization algorithm is a maximum likelihood iterative estimation method for incomplete data. Supposed that \( O \) and \( U \) respectively represent observable and unobservable data. Generally, the parameter vector \( (\omega, \theta) \) of observation data \( O \) can be estimated. Expectation-maximization algorithm includes two parts: (1) calculating the LLF expectation of complete data to \( (O, U) \), where only
data \( O \) is observable; (2) finding the maximized LLF expectation of parameter vector \((\alpha, \beta)\). Above two parts respectively called E-step and M-step, and the process of expectation-maximization algorithm can be represented as mathematical form:

\[
(\alpha, \beta) := \arg \max_{(\alpha, \beta)} E_{(\alpha, \beta)} \left[ LLF(\alpha, \beta) | O, U \right] | O
\]

(5)

Where \( E_{(\alpha, \beta)} \) is the expectation operator of unobservable data \( U \). In order to calculate its LLF, a parameter vector shall be built temporarily. Namely an implicit iterative formula parameter given from Formula (5) updates the parameter until it converges to a given deviation.

B. Parameter Estimation of Reliability Model

Comparing with other model, the matrix exponential calculation is not required when estimating the parameters of reliability model. Suppose that \( Y_i \) is the probability density function of \( i \) th Erlang distribution, its shape and scale parameters, \( \alpha_i \) and \( \beta_i \). Then the form of conditional expectation can be obtained as:

\[
E[Y | F = t] = \frac{p_i p_i(t)}{\sum_{i=1}^{n} p_i p_i(t)}
\]

(10)

then the Formula (9) can be rewritten as:

\[
E\left[ \sum_{i=1}^{n} \tilde{Y}_i | \right] = \frac{u_i p_i F_i(x_i) - F_i(x_i)}{F_i(x_i) - F_i(x_i)} + \alpha \frac{1 - F_i(x_i)}{F_i(x_i) - F_i(x_i)}
\]

(11)

where \( F_i(.) \) is the probability density function of \( i \) th Erlang distribution, its parameters are shape and scale parameters, \( \alpha_i \) and \( \beta_i \). Similarly the expected value of \( E[Y | F = t] \) can be obtained as:

\[
E\left[ Y | F = t \right] = \frac{u_i p_i(t)}{\sum_{i=1}^{n} p_i p_i(t)}
\]

(12)

the E-step process of \( E\left[ \sum_{i=1}^{n} \tilde{Y}_i | \right] \) can be calculated as:

\[
E\left[ \sum_{i=1}^{n} \tilde{Y}_i | \right] = \frac{\sum_{i=1}^{n} \alpha_i p_i \left( F_i(x_i) - F_i(x_i) \right)}{\sum_{i=1}^{n} \beta_i \left( F_i(x_i) - F_i(x_i) \right)} + \alpha \frac{1 - F_i(x_i)}{F_i(x_i) - F_i(x_i)}
\]

(13)

Where \( F_i(.) \) is the probability density function of Erlang distribution, its shape and scale parameters, \( \alpha_i+1 \) and \( \beta_i \), are shown in Formula (10), (11) and (13). The condition and shape parameters of Hyper-Erlang distribution are given in advance, and the computational complexity of above E-step process of Erlang distribution is lesser than the algorithms in similar literatures.
C. Reliability Model Algorithm based on EM

Next, an algorithm is provided by us to determine the shape parameter $a=(a_1,\cdots,a_n)$. The core idea of estimation algorithm is to find an optimal shape parameter $a$, and the total number of phase $\sum a_j$ is given.

As what mentioned before, Hyper-Erlang distribution is the subclass of PH distribution, or rather, all the shape parameters of Hyper-Erlang distribution corresponds to phases. If the total number of phase $\sum a_j$ is given, find a shap parameter to make the process of LLF maximization can be simplified as the emanation of PH distribution, and the scale of phases is $\sum a_j$.

Due to the reliability model based on Hyper-Erlang distribution, the parameters can be estimated by using above EM algorithm process. For the given number of phases $n$, build a vectoring mode set $P_n$ of shape parameter. For example, when $n=4$, the possible mode of shape parameter is:

$$P_4=\{ (1,1,1,1), (1,1,2), (1,1,3), (2,2,2),(4) \}$$ (14)

After the execution of mode recognition, the model with maximum log-likelihood function can be selected. The parameter estimation process of reliability model is shown in Pseudo-code 1.

Pseudo-code 1: EM algorithm of reliability model for generalized grouped data

1. For the given total number of phases $n$, obtain all the modes of Erlang distribution and shape parameter $P_n$.
2. for $n=1$ do
3. Initialize parameters $a$, $r$ and $b$;
4. repeat
5. Using parameters $a$, $r$ and $b$, calculate the expected value according to Formula (20), (23) and (25).
6. update the parameters
   $$\begin{align*}
   a &= \frac{1}{E(\sum_{i}p_{ij})} \\
   r &= \frac{1}{E(\sum_{i}e_{ij})} \\
   b &= \frac{1}{E(\sum_{i}y_{ij})} \\
   \end{align*}$$
7. Meet end conditions.
8. end for
9. Obtain mode $a^*$ on the basis of maximum log-likelihood function.
10. Mode vector $a^*$, estimate the parameters $a^*$, $r^*$ and $b^*$.

IV. EXAMPLE ANALYSIS

The evaluation system of corporate customer relationship ability is constructed by adopting SEM, and the data analysis and processing are conducted by using SPSS 17.0 and Amos 17.0 in the paper. The corporate customer relationship ability is constituted by 6 first-order factor, and has 22 measurement items.

According to Rule 1, the first-order confirmatory factor model has 22 observational variables in total; the number of generating different variances or covariances is $22(22+1)/2=253$. Such model needs to estimate 59 parameters, namely $t=59<253$, which meets the essential condition for model recognition. Moreover, according to Three Indexes Rule: every latent variable of such model shall have three observational variables at least; every observational variable only measures one latent variable; there is no correlation between errors, so the sufficient condition for model recognition is met, and such model can be recognized.

The relevant parameters in model shall be estimated by SPSS 17.0 and Amos 17.0. It is found that the factor loads of A4, C3, F3 and G3 are too low, and shall be deleted. On this basis, a further adjustment shall be made for such model according to the significant results and modified indexes of parameters.

Following analysis is made according to the result of parameter estimation

(1) Reliability analysis. The higher the reliability is, the lesser the influence of error on the score measured from different item in the same scale is. It is studied from two aspects: (1) The $R^2$ value of observed variable reflects the reliability in one factor. Table 1 shows that only the $R^2$ value of B2, C1 and G2 is less than 0.5, the $R^2$ values of other observed variables are more than 0.5. It is shown that single measured reliability of observed variable conforms to the requirement basically. (2) The reliability of first-order factor could be measured by constructed reliability. Table 1 shows that the CR values of six first-order factors are more than 0.6, with excellent internal consistency.

Figure 1. First-order confirmatory factor model
TABLE 1 ANALYSIS ON ADJUSTED FIRST-ORDER CONFIRMATORY FACTOR

<table>
<thead>
<tr>
<th>First-order factor</th>
<th>Measured item</th>
<th>Standardized coefficient</th>
<th>value r</th>
<th>R²</th>
<th>CR</th>
<th>AVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT</td>
<td>A1</td>
<td>0.742</td>
<td>0.551</td>
<td>0.7948</td>
<td>0.564</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A2</td>
<td>0.717</td>
<td>0.925</td>
<td>0.513</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A3</td>
<td>0.792</td>
<td>10.724</td>
<td>0.627</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norm</td>
<td>B1</td>
<td>0.751</td>
<td>0.565</td>
<td>0.7679</td>
<td>0.524</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>0.698</td>
<td>9.652</td>
<td>0.487</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B3</td>
<td>0.723</td>
<td>9.964</td>
<td>0.523</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organization</td>
<td>C1</td>
<td>0.660</td>
<td>0.436</td>
<td>0.8137</td>
<td>0.5227</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>0.749</td>
<td>9.508</td>
<td>0.561</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C4</td>
<td>0.753</td>
<td>9.547</td>
<td>0.567</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C5</td>
<td>0.726</td>
<td>9.275</td>
<td>0.527</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>F1</td>
<td>0.721</td>
<td>0.521</td>
<td>0.7893</td>
<td>0.5559</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F2</td>
<td>0.797</td>
<td>10.237</td>
<td>0.635</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>0.716</td>
<td>9.504</td>
<td>0.512</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooperation</td>
<td>G1</td>
<td>0.796</td>
<td>0.634</td>
<td>0.7073</td>
<td>0.5487</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G2</td>
<td>0.681</td>
<td>8.992</td>
<td>0.463</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coordination</td>
<td>H1</td>
<td>0.716</td>
<td>0.512</td>
<td>0.7688</td>
<td>0.5257</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>0.728</td>
<td>9.611</td>
<td>0.529</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H3</td>
<td>0.731</td>
<td>9.682</td>
<td>0.534</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 2 CORRELATION COEFFICIENT AND STANDARD ERROR OF EACH FIRST-ORDER FACTOR

<table>
<thead>
<tr>
<th></th>
<th>IT</th>
<th>Norm</th>
<th>Organization</th>
<th>Communication</th>
<th>Cooperation</th>
<th>Coordination</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT</td>
<td>1</td>
<td>0.750(0.035)</td>
<td>0.811(0.034)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norm</td>
<td></td>
<td>0.762(0.029)</td>
<td>0.502(0.032)</td>
<td>0.660(0.027)</td>
<td>0.626(0.023)</td>
<td>1</td>
</tr>
<tr>
<td>Organization</td>
<td>0.612(0.022)</td>
<td>0.675(0.027)</td>
<td>0.740(0.022)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>0.657(0.030)</td>
<td>0.656(0.031)</td>
<td>0.728(0.027)</td>
<td>0.683(0.032)</td>
<td>0.853(0.032)</td>
<td>1</td>
</tr>
<tr>
<td>Cooperation</td>
<td>0.661(0.028)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(3) Analysis on goodness of model fit. The Table 1 shows that the specific value $\chi^2/df$ of chi-square value and degrees freedom of model is 1.155, which is less than 2; all the values of GFI, AGFI, NFI, IFI, TLI and GFI are more than 0.9; RMR and RMSEA are less than 0.05. It is thus clear that all the requirements on indexes of fit are met, which means the adjusted first-order confirmatory factor model is effective.

V. CONCLUSION
An analysis method for CRM evaluation based on parameter estimation of expectation-maximization reliability model is proposed in this paper. It builds the evaluation system of corporate customer relationship ability by SEM, and analyzes the evaluation through the parameter estimation of expectation-maximization reliability model proposed. The empirical result shows that each evaluation index of corporate customer relationship ability is interactional rather than independent. Among those, the support ability of organization system has the highest standardized coefficient to the customer relationship ability, which has shown the core roles of organization and personnel in the CRM.

REFERENCES
Research in Education, 2014, 3(3).

foreign direct investments: A MNC headquarters challenge[J].

variation of out-migration rates[J]. Demography, 1988, 25(4):567-
579.

[12] López-Nicolás C, Ángel L. Meroño-Cerdán. Strategic
knowledge management, innovation and performance[J].
International Journal of Information Management, 2011, 31(6):502-
509.